

Published in final edited form as:

Trends Immunol. 2014 May ; 35(5): 195–204. doi:10.1016/j.it.2014.03.001.

The Chromatin Landscape and Transcription Factors in T-Cell Programming

Ellen V. Rothenberg

Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125 USA

Ellen V. Rothenberg: evroth@its.caltech.edu

Abstract

T-cell development from multipotent progenitors to specialized effector subsets of mature T cells is guided by the iterative action of transcription factors. At each stage, not only do transcription factors interact with an existing landscape of histone modifications and nucleosome packing, but they also interact with other bound factors and modify the landscape for later-arriving factors, in ways that fundamentally affect the control of gene expression. This review covers insights from genome-wide analyses of transcription factor binding and resulting chromatin conformation changes that reveal roles of cytokine signaling in effector T-cell programming, the ways one factor can completely transform the impacts of previously bound factors, and the ways that the baseline chromatin landscape is established during early T-cell lineage commitment.

Keywords

Cis-regulatory element; T-cell development; Histone modification; CD4⁺ T-cell subsets; Genomics

Complex developmental processes and transcription factor guidance

Lymphocytes develop from multipotent stem cells through a hierarchical process in which the spectrum of accessible cell fates is progressively narrowed, even as the cells become armed with an increasing kit of expressed effector genes (reviewed in [1–6]). This differentiation program is actually the output of four kinds of processes, operating repeatedly at a series of choice points: (1) new regulatory gene activation, (2) selective earlier regulatory gene silencing, (3) combinatorial use of previously expressed and newly expressed regulatory factors across the genome, and (4) changing patterns of accessibility of cis-regulatory elements for transcription factor binding in general, as a result of past and ongoing regulatory events. These interlaced gene regulation changes underlie the whole

© 2014 Elsevier Ltd. All rights reserved.

Corresponding author: Rothenberg, E. V. (evroth@its.caltech.edu).

Conflict of interest statement

The author is not aware of any conflict of interest that could inappropriately influence this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

history of any given CD4⁺ effector T-cell from its origin in totipotent embryonic cells, through specification as a hematopoietic stem cell that can give rise to all possible blood cell types, eventually through a committed T-cell precursor that can still generate multiple classes of T cells, through a choice between helper and killer fates in the thymus, and finally through a post-thymic choice of alternative effector fates[7–14]. By the end, the cell is finally primed for expression of the right combination of cytokines and signal receptors to play its appropriate role in the immune response, although this last effector choice remains somewhat plastic under variable physiological conditions[10, 15–17]. However, all the earlier choices that establish a T-cell identity are durable. The way that present gene regulatory mechanisms set up opportunities for future gene expression while removing the conditions needed for past gene expression gives the program its hierarchical, one-way character. With the advent of new technologies for visualizing the interaction of transcription factors with DNA and the local consequences of that binding for regulatory sequence activity, a wealth of new insights into this complex process have emerged. This review looks at exciting implications of recent discoveries for the prospect of understanding lymphocyte development as a whole at a causal, molecular level.

Sequence-specific DNA recognition gives transcription factors the primary responsibility as readers of the regulatory genetic code, and this is particularly important in any developmental process. If a cell needs to change its gene expression pattern, it must use the combined activities of the transcription factors it expresses to mark out the right new regions of the genome for transcriptional activation or to signal specific, previously active regions to become inactive. This is not a role that can be substituted by histone modification reader/writers that simply propagate a chromatin state irrespective of sequence, nor by passively replicated CpG methylation patterns. At each stage in a developmental process, as new transcription factors are expressed, they can bind to previously empty regulatory sites with consequences that affect the activity of cis-regulatory elements like promoters and enhancers[18–23].

A valuable shortcut toward understanding lymphocyte development has been the identification of a few transcription factors that have potent, essential functions in the emergence of particular cell types. They achieve their impact not only through direct regulation of effector genes but also through positive and negative cross-regulation of collaborating and antagonistic factors, respectively, in a gene regulatory network. EBF1 and Pax5 are powerful regulators of B-cell development, Th-POK (Zbtb7b) is a powerful regulator of CD4⁺ T cell development, whereas for effector subtype specialization of CD4⁺ T cells, T-bet, GATA-3, ROR γ t and Foxp3 are the signature regulators of Th1, Th2, Th17, and Treg cells, respectively (rev. in [9, 24–26]).

Analogously PU.1 and C/EBP family factors dominate myeloid cell development[27]. These central lineage-regulatory factors have attracted much attention because of their influence, and are often referred to as “master” regulators (see Box 1). In detail, however, the causal links through which any transcription factor brings about a transformation of cellular function depend on system aspects besides the transcription factor’s own ability to recognize its target sites in genomic sequence. The choice of sites to bind, and the possible effects exerted when the factor binds there, are dependent both on the current accompanying

presence of other transcription factors and on the history of the cell's prior regulatory states. Understanding what actually happens at a developmental choice is more interesting than a simple regulatory command.

Box 1

Beyond the “Master regulator” concept

The roles of EBF1 and Pax5 in B cell development provided a strong prototype for powerful central lineage regulators (reviewed by [26]). These are both specific for B cells, expressed only in B-lineage cells among all hematopoietic lineages; they are both essential for B cell development; they both directly activate B-cell genes; and they can both antagonize development to non-B cell fates. By simply focusing on the requirements for activating these two transcription factors and the activities of these two factors on their targets, an immense amount of insight into B-cell development has been gained. However, at the mechanistic level, EBF1 and Pax5 work collaboratively with other transcription factors such as E2A, which is not B-cell specific. The same requirement for combinatoriality tempers the roles of all “master regulator” candidates[9, 24, 25].

This is particularly true for the CD4⁺ T cell subset regulatory factors, GATA-3, T-bet, ROR γ t, and Foxp3. In each cell type these factors must collaborate with STAT or Smad transcription factors activated by appropriate cytokine receptor stimulation in order to work. Furthermore, despite their lineage-specific roles in late T-cell effector differentiation, expression of GATA-3, T-bet (*Tbx21*), and ROR γ t (*Rorc*) is not specific to the CD4⁺ T-cell lineage programs in which they predominate. T-bet is a prominent regulator in NK cells, iNKT cells, and CD8⁺ T cells as well as Th1 CD4⁺ cells, where at least its roles as a positive regulator for IFN γ expression may be conserved. However, GATA-3 and ROR γ t have crucial earlier stages of expression that affect the fate of all $\alpha\beta$ T-cell precursors, clearly regulating different sets of genes than the Th2 and Th17 programs. Turned on in the first wave of T-cell gene expression induced in the thymus, GATA-3 is essential for the initial specification and survival of thymocytes in the earliest T-cell developmental stages, then again for β -selection, and again for choosing the CD4⁺ and not the CD8⁺ cell fate during positive selection[85]. It continues to be important in CD8⁺ cells, albeit at a lower level than in CD4⁺ cells[86]. ROR γ t turns on as the cells pass β -selection into the CD4⁺ CD8⁺ (DP) stage, where it is crucial for maintaining survival of cells awaiting any form of positive selection at all, long before effector lineage is determined[87]. The gene expression patterns that depend on GATA-3 and ROR γ t in these earlier stages have little to do with the gene expression patterns that they govern in late effector T-cell development[39, 85]. Thus, developmental context dictates the range of action that is open to these factors.

Transcription factor binding vs. transcriptional regulatory activity

Genome-wide analysis of transcription factor binding using chromatin immune precipitation and deep sequencing (ChIP-seq) has revealed some unexpected general aspects of these factors' function. The numbers of genomic sites detectably bound by many transcription factors greatly exceed the numbers of genes that are shown to be affected by gain or loss of

the factors' activities. Also, although binding sites are often clustered around genes that the factor clearly regulates, they are also commonly found associated with non-responsive genes. This means that binding (e.g. detected by ChIP-seq) cannot simply be assumed to be functional. In fact, patterns of local occupancy can fail to predict function in both directions. The "false negative" problem is that one can readily find genes that are affected by changes in level of a transcription factor even though that factor appears to lack any binding site near the gene's promoter. The "false positive" problem is that many factors bind to a range of genomic sites at which the nearest genes do not change expression in response to the factor.

The "false negative" problem can be explained by two kinds of mechanisms that are each intrinsically interesting: an indirect gene network connection, or a distal cis-regulatory element. If a regulated gene actually depends on an intermediate regulatory or signaling gene that is truly a direct target of the factor of interest, then the regulated gene can be affected by gain or loss of function of the factor of interest even if it is not a target itself. This is not biologically insignificant but a clue to a major feature of developmental control[28]. Gene-specific indirect effects, once dissected by epistasis tests, reveal large sectors of gene network architecture.

In another source of false negative results, an affected gene apparently lacking a transcription factor binding site can still be a direct target, if that gene has a distant cis-regulatory element that lies outside of the search space assumed to belong to that gene. Even if binding is not found in a region thought to be a gene's enhancer, increasing evidence suggests that genes can possess multiple potential distal elements rather than a single enhancer. Some of these elements may simply appear to reinforce each other's activity in the cell types under consideration ("shadow" enhancers) while others are functionally quite distinct[29]. Cis-regulatory elements for *Gata3* and *Bcl11b* in mouse T-lineage cells have recently been found in gene deserts 280 kb and 850 kb downstream of the genes, respectively[30, 31]. Thus, extensive searches far beyond the 10-kb range, coupled with strategies for evaluating function, are increasingly important for understanding of mammalian gene regulation. To map promoter/enhancer interactions on such a large scale, multiple new techniques have been developed to map long-range promoter/enhancer looping based on crosslinking and deep sequencing[32–35], and these are beginning to provide considerable information about the organization of active and inactive genes and their regulatory elements in the nucleus[34, 35].

Many "false positive" signals really do not indicate function. A factor that engages in strong interactions with other factors may join a bound factor ensemble even though a full activating quorum has already been established without it (see below). Multiple-occupancy regions, especially some extensive ones called "superenhancers"[36], are highly likely to be important cis-regulatory elements[37–39], but it is not clear how many of the factors involved in binding at such regions are really required for activity. When a factor is bound at such a site, it may simply act as a marker for an active cis-regulatory element without being an essential contributor. At the opposite extreme, certain factors that have the power to bind directly to nucleosome-occupied DNA need not bind where other factors are engaged[21]. In some cases they can establish occupancy at isolated sites in chromatin that have no potential for functional activity. However, other suspected false positives are a consequence of the

way that transcription factor action interfaces with the chromatin regulatory state to affect future action of other transcription factors at the site. A closer examination of the steps toward cis-regulatory element activation suggests that factors can play certain roles through isolated binding to nucleosome-packed DNA which may become important for later transcriptional regulation, even if they do not correlate immediately with target gene expression.

Target seeking: constraints of context and history

Factors like EBF1 and GATA-3 occupy ~1500–4000 sites in pro-B and -T cells, SCL(Tal1) and Runx1 bind ~10,000–15,000 sites in early hematopoietic progenitors, whereas PU.1 and Pax5 bind ~30,000–75,000 depending on cell type[18, 37, 38, 40–48]. Yet even those that bind to many more sites reach only a fraction of the full spectrum of their potential sites in the genome as defined by analysis of DNA sequence. A transcription factor is often able to bind particular sites in one cell-type context that it cannot bind at all in another. The sites bound by PU.1 are substantially different in macrophages than in B cells and in early T cells[42, 43]. Pax5, which binds to a large number of sites in B-lineage cell genomes, binds to a different spectrum of targets in pre- and pro-B cells than in mature B cells[47]. GATA-3, despite its reputation as a “master” regulator of the Th2 cell fate, is recruited to very different genomic sites in developing T cells than in Th2 cells[41]. It even redeploys to distinct patterns of occupancy during intrathymic development, between the early T-cell developmental stages (ETP-DN2b) and the intermediate CD4⁺ CD8⁺ (DP) pre-selection stage thymocytes[41, 42](Fig. 1). These redistributions are necessarily guided by partner factors and/or underlying chromatin landscapes.

Sites for some transcription factors can be effectively occluded based on a cell’s developmental history. A good example is the fate of exogenously transfected EBF1 in EBF-negative hematopoietic and non-hematopoietic cell types[40]. EBF1 normally regulates one set of target genes in B cells and another, completely distinct set in adipose (non-hematopoietic) cells. When introduced into EBF-negative hematopoietic cells, exogenously transfected EBF1 becomes bound to “B-cell gene” target sites where it induces activating histone modifications. However, it does not reach these B-cell target sites at all if transfected into EBF-negative non-hematopoietic cells[40]. Thus, earlier hematopoietic developmental events must set a stage in which EBF1 can work to promote B-cell gene expression. Similarly, exogenous PU.1 introduced into T-lineage cells occupies fewer sites than it naturally occupies in myeloid cells[49]. In some cases the differential accessibility may reflect the presence or absence of additional partner transcription factors, which allow otherwise transient, undetectable interactions of the factor with DNA to be stabilized at particular sites. In other cases, differential accessibility can reflect a barrier of closed chromatin. In the earliest T-lineage cells, which do express PU.1 normally, PU.1 fails to bind to sites with the repressive chromatin mark, H3K27me3, even when they are known sites of PU.1 action in other cells[42]. On the other hand, a factor can be recruited to a new spectrum of sites by the presence of other factors. Expression of factors such as EBF1, Runx1, PU.1, or STAT factors in certain contexts can cause the redeployment of other, previously expressed regulatory factors, to join them in multi-factor complexes at new sites[18–20, 43, 48].

Cis-regulatory elements enable transcription factors to act combinatorially across time

Partner effects on transcription factor binding are in complete harmony with extensive evidence that effects on gene expression are determined by combinations of transcription factors, not by factors acting alone. Cis-regulatory elements, which represent clusters of potential transcription factor binding sites, act as a scaffold to dictate the combinations of transcription factors that will be needed to cross the threshold for exerting activity. Transcription factor collaboration at these sites can be based on strong cooperativity for DNA binding, where interactions with a partner factor can deeply influence a transcription factor's stability of DNA binding[50–54]. On some cis-regulatory elements, coordinate occupancy by multiple factors appears to be required in order to assemble a stable ensemble of binding factors at all. Such “enhanceosome” assembly is an elegant mechanism to ensure “AND” logic for physiologically high-stakes gene activation pathways such as induction of some cytokine genes[55]. However, functional cooperativity can also emerge at cis-regulatory elements where transcription factors bind separately in time. For example, one factor can bind initially as a “pioneer” and await the advent of others that eventually complete the quorum for enhancer or promoter activity[22, 24, 48, 56, 57]. Thus, without a needed partner, a factor may be unable to make a particular cis-regulatory element work even when it does bind. Considerable recent work in several model systems has now defined component steps that go into activating a cis-regulatory element, by combining evidence from transcription factor binding analysis (ChIP-seq) with analyses of histone modifications (ChIP-seq) and general DNA accessibility (DNase-seq, FAIRE, or restriction enzyme accessibility assays)[21, 22, 58]. This process is summarized in Fig. 2.

For most of the genes used in lymphocyte development, the onset of expression involves activation from a silent ground state (see Box 2). Generally the most cell-type specific, developmentally dynamic chromatin modifications take place at enhancers (distal cis-regulatory elements)(e.g. [42, 43]). Long before a gene's expression the promoter may already be marked by a histone “accessibility” mark (H3K4me2 or even H3K4me3, emplaced by MLL histone methyltransferase factors)[42, 59, 60], although in these non-expressing cells, if the promoter itself has CpG islands, it may initially be occluded with repressive H3K27me3 histone marks (H3K27me3; see below). Such repression marks persist by default but can readily be removed by transcription factor-recruited cofactors when a gene is activated[42, 61]. At the ground state for enhancers, the DNA is often unmarked but relatively condensed in nucleosomes, as defined by relatively poor DNase sensitivity. The opening and activation process at the enhancer begins when certain transcription factors are expressed, which are capable of binding to these nucleosome-wrapped DNA sequences without impediment[58]. The binding of these “tier A” or “pioneer” factors initially need not lead to any more detectable consequences, but it may create a preferential recruitment site for other, “tier B” factors as they become available.

Box 2

A complex regulatory baseline in hematopoietic precursors

Many of the genes activated in the T and B cell programs are expressed in a highly cell type-specific way, beginning after the divergence of distinct types of partially restricted hematopoietic progenitors. T-cell and B-cell regulatory genes like *Bcl11b* and *Pax5* are embedded in repressively H3K27me3-marked chromatin in hematopoietic progenitors until they are activated[31, 42, 59, 60], and *Bcl11b* is extensively methylated across its promoter and enhancer regions until T-cell development is under way[77]. These cases suggest genuine epigenetic inaccessibility. The closed state of these genes may continue or be reinforced in developmental lineages that do not turn them on, explaining why *Bcl11b* in B-lineage cells and *Ebfl* in T-lineage cells are sequestered in repressive intranuclear domains[32]. However, the silence of these loci in multipotent progenitors is obviously not permanent. This suggests that a tier A-type activity must be induced in early specification of T and B lineage cells to begin activation of the appropriate genes. Alternatively, combinations of factors that can collaborate in tier A/tier B fashion may need to rise above some threshold concentrations, simultaneously within the same cells, in order to be able to open up the regulatory elements of these genes.

In fact, there is indirect evidence that multiple factors that activate lineage-specific genes are already present in hematopoietic precursors, even if they are not yet strongly or stably enough expressed to coordinate an entire lymphocyte developmental program. Individual uncommitted hematopoietic precursors are known to express low-level transcripts from genes associated with several distinct lineages at once[88–91]. This phenomenon, multilineage priming, does not embrace all the genes that will later be expressed as part of the final differentiation programs; many remain silent. However, the genes being co-expressed belong to the same developmental programs that these precursors may access, implying that some transcription factors that could drive these programs are already present in sufficient concentration to promote at least intermittent target transcription.

What makes multilineage priming-associated genes more accessible than *Bcl11b* and *Pax5*? Some lineage-associated genes may initially reside in open chromatin, awaiting a positive regulator. Many have accessible marks (H3K4me3) at their promoters even in hematopoietic stem cells[59]. More tightly regulated genes may require a lineage-specific tier A factor. Another difference may reflect the constraints of transcription factor collaboration across time: how strictly tier A/ tier B factor activities must be coordinated to open an enhancer element. Recent single-cell gene expression analyses show that hematopoietic multilineage progenitors are heterogeneous in their expression of key transcription factors[92–94]. The range of phenotypes suggests that transcription factor expression could vary dynamically in these precursors even when the same general combinations are present, promoting the same fates (cf. [95]). Genes with cis-elements where tier A factors can wait for tier B recruitment can easily be turned on. Genes with elements requiring coincident binding by two or more tier A/B factors may not be effectively activated until the gene regulatory network that mediates differentiation stabilizes expression of these factors at adequate levels.

Either tier B factors or the initial tier A factors, depending on the factors and the cis-regulatory element, can then begin opening the nucleosome packing around the binding sites[58, 62]. As nucleosomes are pushed apart, the element can become DNase hypersensitive, and usually the histones are modified to H3K4me1 or further modified to H3K4me2, using MLL histone methyltransferases[18, 19, 43, 44, 63](Fig. 2). This is a critical tipping point in the progression to enhancer activity, but even so, it can occur long before the enhancer is actually active. Note that tier B factors themselves need not be expressed after tier A factors; they can actually pre-exist in the cell. However, they do not come to these sites until the tier A factor creates a beachhead, and this is detected as recruitment or redeployment of the tier B factor. Examples include the impact of EBF1 expression on E2A binding, or the impact of Runx1 expression on SCL/Tal1 binding[19, 48]. These initial binding events can occur coordinately, in which case the cooperative binding of two classes of factors enhances the binding of both. But the tier A/tier B distinction is worth making because in other cases, single occupancy can exist for a long time before the recruitment of second-tier factors can trigger full enhancer accessibility[18].

The modified histones induced by tier A/B factor binding indeed appear to mark “poised enhancers”, sites of long-term preferential accessibility[18, 64]. To get from this point to positive regulatory activity of the enhancer, however, usually requires binding of yet more factors (or possibly removal of obstructive ones). With “tier C” factor addition, then, the element suddenly emerges as an “active” enhancer[24, 44, 65], characteristically marked with H3K27Ac. This results as the assembly of transcription factors achieves a quorum for recruitment of the EP300 (p300) or Crebbp (CBP) transcriptional coactivator complexes, which acetylate local nucleosomes at the H3K27 sites[65, 66], in what may be a hit-and-run process[63]. This is well coordinated with enhancer looping to a cognate promoter via the Mediator complex, and the target gene finally responds to the accumulation of events that have activated its enhancer.

Because binding of different factors is so affected by binding of prior factors in this process, the transcription factor that appears to correlate best with the activity of the target gene depends on which type of factor is limiting in the cell. This is not a biochemical question so much as a developmental programming question. In fact, a given factor can play different roles at different cis-regulatory elements. For example, in myeloid cells, development is dependent on C/EBP factors and PU.1, and these can each act as initial pioneering factors for the other’s recruitment at different sites[43]. But the actual transcription onset depends on the addition of the tier C factor(s) to the tier A/tier B engaged complex. Thus, if the tier C factor(s) is rate limiting, and the binding of the tier B factor(s) is always saturated, tier C may look like the “master regulator”. But if the tier B factor(s) is not present, then there may be no open chromatin at the enhancer for the tier C factor(s) to bind to, and thus a factor playing the tier B role may appear more dominant.

An asymmetry of understanding: the obscurity of repression

Cis-regulatory elements also include the sites where repression complexes are assembled. Repression is less well understood in general. However, most transcription factors can be bifunctional, participating in ensembles that recruit coactivators or corepressors depending

on the overall architecture of transcription factors organized by their DNA binding at a given cis-regulatory element. At present, there is only sparse evidence about how many copies of a repressor or how many different kinds of binding factors at a given site are needed to establish repression. Converting an active cis-regulatory element to a repressive one usually involves local expulsion (or cell-wide loss) of positive regulators, with loss of histone acetylation and often reduced H3K4 methylation as well[46, 67]. These elements also commonly become less DNase-accessible as the repression process compacts the nucleosomal array[68]. Unfortunately for systematic analysis, these elements thus become less detectable against the background of neutral DNA. Some repressed elements become modified by Polycomb Repressive Complex 2 to add H3K27me3 marks, especially elements near repressed CpG-island promoters[67, 69, 70]. But this process is often a slow step that occurs long after transcription of the repressed gene has been shut off[42], and therefore gives only indirect evidence of the factors that caused repression and the critical sites through which they worked. A better understanding of repression is crucial for explaining commitment (Box 3).

Box 3

Two problems for the future

- Enhancer-promoter interaction rules need elucidation. Enhancers can be detected readily when changes in their histone marks correlate dynamically with those of a neighboring gene[31, 65], but this matching is extremely difficult when multiple genes lie between the element and the promoter it actually regulates. Still to be defined is how specific enhancers and their “cognate” promoters identify each other across distances that can reach multiples of 100 kb of DNA; how this traffic is regulated; and how many distinct enhancer elements are actually communicating with a promoter at any given time to control its activity level. New crosslinking and deep sequencing variants of chromosome conformation capture technology like 4-C, 5-C, hi-C and ChIA-PET offer some promise for addressing the mapping part of this question[32–35]. At the same time, insight into the temporal dynamics of enhancer-promoter communication may begin to emerge through live fluorescent monitoring of single-cell gene expression level fluctuations within a developmentally equivalent population. Determining how much of gene expression is steady and how much results from “bursts” alternating with silence may help to establish some boundary values for the stability or instability of promoter-enhancer associations. The technologies that must converge to answer this question are in rapid development.
- Repression is central to the establishment of a coherent developmental identity, and it plays an even more prominent role in the pathway toward T-cell development than might have been expected[42]. However, at a genome-wide level the mechanisms used for different kinds of repression are poorly defined: in particular, the distinction between temporary “repression” and permanent “silencing” is currently impossible to recognize. DNA methylation, at least as detected with current methodology, does not appear to be a barrier to gene activation[96] and is easily removed, e.g. in cases of the T-cell regulatory genes

Tcf7 and *Bcl11b* which are demethylated at multiple sites as they become active[77]. Despite the usefulness of H3K27me3 to denote one form of repression, genes with H3K27me3 marks need not be profoundly repressed. For example, the CD4⁺ specification factor *Zbtb7b* (Th-POK) is active in the earliest T-cell precursors but silenced with H3K27me3 marking during commitment [42], only to be reactivated a few stages later during positive selection[97]. It can be difficult to locate the DNA target sites where critical repressive interactions are taking place, since in many cases, repression leads to a disappearance of transcription factor binding as well as loss of active chromatin marks and DNase accessibility[47, 67]. This can be a consequence of chromatin condensation machinery, possibly combined with intranuclear compartmentalization[32, 84, 98]. To improve understanding of these mechanisms, repression may be best dissected by taking advantage of the developmental kinetics of the regulatory change, to catch factors in the act of imposing the repressed state.

Within the cis-regulatory module: strict ordering, late arrival, strong impact

If most active enhancers are sites not for binding of single factors but for binding of large ensembles of factors[37–39], it is easy to see how their collective impact to displace nucleosomes could establish a “virtuous circle” for enhancer activity. At the same time, however, this group action can obscure the logic of what each factor is truly adding to the activity of the cis-regulatory element. Thus, it has been illuminating to study two elegantly documented cases in mature CD4⁺ T cells where it is clear that collaborating transcription factors act in a particular, strict temporal relationship in order to activate or repress their target genes correctly. These factors clearly do not act on the tier A level, but as quorum-completing tier C or late tier B factors that associate with already DNase hypersensitive, H3K4me1-marked regions to determine the output regulatory function of these elements.

Conventionally, Th1 and Th2 cells are thought to have their function determined by T-bet and GATA-3, respectively, sometimes called their “master” regulators (Box 1)[9, 24, 25]. In fact, input from IL-12 via STAT4 (or IFN γ via STAT1) is also crucial for Th1 development, and input from IL-4 via STAT6 is needed for Th2 cell development. The respective STATs, which translocate to the nucleus only when these cytokine signals are present, are known to provide viability or activation maintenance in a general sense. They also play roles in upregulating T-bet or GATA-3 initially, but their functions are often imagined in supportive roles only. Now, however, new evidence demonstrates that the STAT factors act much more pervasively to control the entire program of gene expression in their respective T helper subsets[71]. Despite sustained differences in T-bet and GATA-3 expression between Th1 and Th2 differentiated cells, there are minimal differences between them in terms of the H3K4me1 modification states of sequences across the genome. Regions that will become both Th1-specific and Th2-specific gene enhancers are found equally marked by tier A/B processes in both cell types. However, conversion of any of these Th1 or Th2-specific enhancers to “active” status by recruitment of EP300 depends on the further binding of appropriate STAT factors (STAT4 or STAT6) in response to current cytokine signaling.

This is the step that determines the activity of the neighboring/associated/cognate gene. Lineage-inappropriate genes also lose EP300 from their enhancers in a STAT dependent way.

The structural requirements for enhancer activation at many of these cis-regulatory regions reveal hierarchical roles of different transcription factors in successive stages of enhancer activity regulation. A major lesson in this work is that STAT4 and STAT6 have a much larger impact on recruiting EP300 than the “lineage determining factors” T-bet and GATA-3[71]. T-bet and GATA-3 recognition motifs are found enriched in the same Th1 and Th2-specific enhancers that are activated by STATs, although unfortunately it is not clear whether T-bet and GATA-3 binding is actually stable without STAT co-occupancy at these sites. Because STAT activity is needed to turn on T-bet and GATA-3 during Th1 and Th2 differentiation, the STAT mutants themselves could suffer from compound defects. However, the relative roles of these factors have been revealed in powerful add-back experiments to complement the STAT knockout defects. For most of the Th1- or Th2-specific enhancers, the requirement for STAT activation to induce EP300 recruitment to these H3K4me1-marked elements cannot be bypassed by forced expression of T-bet and GATA-3. Therefore the sequences of these Th subset-specific enhancers enforce the requirement for current cytokine signaling, through engagement of the correct STAT, in order to trigger target gene expression.

A second example reveals the opportunistic *modus operandi* of the transcription factor Foxp3, which is often considered the “master” regulator for Treg cells[72, 73]. Quite unexpectedly for such a potent factor, when its binding distribution was evaluated, it turned out that Foxp3 cannot find its own target sites in the genome. It depends completely on the presence of other factors at an enhancer to enable its own binding and function and appears completely unable to get to DNA sites by itself[74]. Fortunately, Foxp3 is not required for viability of naive T cells, and so cis-regulatory element activity measured by DNase hypersensitivity can readily be compared between Foxp3⁺ induced Treg cells and Foxp3-negative cells that could represent their precursors. This analysis shows that all the sites Foxp3 binds to are already DNase hypersensitive in Foxp3-negative cells. Ultra-deep sequence analysis of the DNase hypersensitivity patterns in these samples has made it possible to map individual transcription factor contacts within active enhancers in the presence and absence of Foxp3. In the fine structure of DNase protections within the core of the DNase hypersensitive regions, each 10–15 bp feature represents a kind of protective shadow diagnostic of transcription factor binding at that particular site[75, 76]. This has revealed not only that Foxp3 binding causes no change in DNase hypersensitivity, but also has identified the exact motifs within those cis-regulatory regions that are bound by transcription factors with and without Foxp3[74]. Foxp3 binding site selection to the genome does not involve recognition of its own cognate motifs, but rather, sites where other factors are already bound in an unchanging way. What Foxp3 affects is the functional consequence of these transcription factors’ binding.

Two classes of these sites contrast in terms of the likely order of action of the site-determining factors and Foxp3. The most common sites are engaged by Runx and Ets family transcription factors. Members of these families should be active in all T cells and could be

playing classic tier A/B roles. However, other Foxp3 recruitment sites include sites for NFAT, a factor that only enters the nucleus when Ca^{2+} release activated by T-cell receptor (TCR) signaling causes its translocation (somewhat analogously to STAT activation by cytokine receptors). In this case, Foxp3 already present in the nucleus presumably piggybacks onto newly immigrated NFAT during TCR stimulation. This brings Foxp3 to sites where NFAT could normally mediate Ca^{2+} dependent gene activation, which Foxp3 then blocks. The only places where Foxp3 binds to Forkhead sites that ought to be its cognates are where it can displace a fellow family member, FoxO1, which had been acting as a tier A/B pioneer factor. These again are preferentially sites where Foxp3 represses, i.e. by reducing the positive activity that the FoxO1 had formerly provided[74]. Thus, despite its potency at transforming T-cell function, Foxp3 cannot dictate cellular properties through its own DNA binding specificity, but only modify them within boundaries established by the factors that determine that cell's open enhancer landscape.

Creating the ground state for lymphocyte identity

Both of these examples reflect the importance of the prior enhancer accessibility landscape, but only hint at how it is established. The pre-marking mechanism has been studied very little to date because it needs to be studied longitudinally across a developmental process, focusing not only on the cells of ultimate interest but also on purified cells in successive precursor stages. Early T cell development provides a rare opportunity to catch sight of the process through which the regulatory landscape around many lymphoid development genes is changed from a progenitor-like state to a lineage-specified T-cell state. Because of the fine-scale dissection of stages that is possible in this system, some initial characterization is now in hand[42]. T-cell precursors transition from multipotency to commitment from the Early T cell Precursor (ETP) to the DN2b/DN3a stages (Fig. 1), before the TCR genes become rearranged and thus before any TCR-dependent selection[3]. During commitment, there are dynamic histone marking changes across multiple genetic loci[42, 60], coupled with major transformations of DNA methylation status at key sites[77]. Although the specific roles of individual transcription factors in this process are still being defined, the “crime scenes” of multiple distinct regulatory events can be mapped clearly. From stage to stage, specific cis-regulatory elements change from latent (or silent) to active modification states, while new repressive marks are deposited on sites that were recently marked as enhancers, after the positive inputs have been lost[42, 60, 78]. By following H3K4me2 across the genome from stage to stage, it is possible to detect first appearance of specific marked distal elements, presumably at the tier B stage, that immediately foreshadow the activation of linked, key developmental genes. These marking changes are functionally significant: for example, removal of H3K27me3 and appearance of H3K4me2 at a previously unknown, distal region in parallel with those at the promoter of the tightly regulated *Bcl11b* gene has revealed a novel, far-distant enhancer for this gene[31]. Such results in aggregate should provide initial guidance to help locate many developmentally significant cis-regulatory elements for T-lineage cells.

Like many hematopoietic cells, T-lineage cells preferentially use enhancers that are rich in sites for Ets-family and Runx family transcription factors regardless of what other factors bind there. This is not only the case for Tregs. In Th1 and Th2 cells, most active enhancers

whether shared or subset-specific are enriched for Runx motifs[71]. Basic helix-loop-helix (bHLH) factors like E2A and HEB tend to bind to Ets or Runx motif enriched sites in B and T cell precursors[19], and an Ets family motif was as enriched as the GATA motif in a sample of GATA-3 binding sites in early T cells[42]. These families of factors are indeed expressed throughout T cell development, but they cannot be simply responsible for defining T-lineage specific sites since their roles go back to multipotent hematopoietic progenitors. Multi-factor ChIP-seq studies of primitive multipotent hematopoietic progenitors have repeatedly shown enhancers for the signature genes of these cells to be marked by the clustering of bHLH, GATA, Ets, and Runx family factor binding. T lineage cells share with hematopoietic precursors expression of factors of all of these families. Although individual Ets-family members change expression, the Ets1/Ets2/Fli1/Erg subfamily (with nearly identical binding specificities) is consistently represented by at least one member, from stem cells to mature T cells[79, 80]. Thus, such Ets and Runx factors are likely to play a tier A hematopoietic role, but not a role to define the lineage-specifically poised status of developmentally emergent enhancers.

What appears unexpectedly prominent during lineage commitment is extinction of previously expressed regulators. As the cells lose multipotentiality, the regulatory genes that were active in the initial T-cell development stages are shut off and then silenced[42, 79–81]. The cells lavish considerable regulatory complexity on this task, imposing repression on different genes in at least three waves from ETP stage to the β -selection checkpoint, and utilizing a variety of different repressive mechanisms[42]. PU.1, the abundant factor commonly playing tier A/B roles in progenitors, is one of the factors eliminated during commitment. As the cells cross the β -selection boundary and develop to the CD4⁺ CD8⁺ (DP) stage, there is a significant deposition of H3K27me3 marking on promoters of numerous regulatory genes that were active in initial stages, even though the silencing of transcription and loss of positive enhancer inputs had occurred earlier[42, 60]. By the time the cells reach DP stage, the progenitor-associated regulatory genes that have CpG island promoters are already fully silenced with H3K27me3 marks[42, 82, 83]. The loss of these foundational regulators not only removes factors that might recognize specific DNA motifs, thus closing off certain elements that required their occupancy, but also may shift the status of the transcription factor complexes that remain. For example, although the crucial bHLH factor E2A remains present throughout, its involvement probably shifts from membership in SCL (Tal1)/E2A or Lyl1/E2A heterodimers to functionally different E2A/HEB and E2A/E2A homodimers as first *Tal1*, then *Lyl1* are repressed. Thus, certain accessibility barriers are already in place, a large part of establishing what the later regulatory state of mature naive T cells will and will not allow.

Concluding Remarks and Outlook

Transcription factors have great power over cell function through gene regulation, but they cannot operate dictatorially outside of a regulatory state context. They must work combinatorially with other factors, whether the contributions of those other factors are simultaneous or not. The ability of transcription factors to modify nucleosomal packing and recruit local histone modifications gives them a way to propagate their influence across time. This is particularly important in a hierarchical, prolonged developmental process like

hematopoiesis which re-uses similar transcription factors repeatedly to activate divergent gene programs. Fortunately, we can now use some of these chromatin modifications to detect when prior action of a transcription factor ensemble has altered the thresholds for local action by new factors. By considering the multiple tiers of function that a transcription factor can provide at a given cis-regulatory element, it is possible to understand how gene expression programs that differ intricately between lymphocyte subsets can be established by factors whose individual expression patterns and even whose binding patterns appear to be too broad.

The gene networks that control T-cell development are much deeper than a single tier of gene regulation, and the crucial frontier for the future will be to use these new tools to understand how the regulators themselves are controlled at successive stages. This will involve more elucidation of distant cis-regulatory region interaction with target genes, of the rules for positioning genes in permissive and repressive nuclear compartments[84], and a much clearer understanding of repression. However, the pathway to understanding this process is becoming clear and a causal, molecular mechanism for the levels of T-cell development is coming within reach.

Acknowledgments

The development of these ideas was supported by NIH grants CA148278, CA90233, AI095943 and HD076915, and the Albert Billings Ruddock Professorship.

REFERENCES CITED

1. Yang Q, et al. T-cell lineage determination. *Immunol Rev.* 2010; 238:12–22. [PubMed: 20969581]
2. Mercer EM, et al. Factors and networks that underpin early hematopoiesis. *Semin Immunol.* 2011; 23:317–325. [PubMed: 21930392]
3. Rothenberg EV. T cell lineage commitment: identity and renunciation. *J Immunol.* 2011; 186:6649–6655. [PubMed: 21646301]
4. Luc S, et al. The earliest thymic T cell progenitors sustain B cell and myeloid lineage potential. *Nat Immunol.* 2012; 13:412–419. [PubMed: 22344248]
5. Welinder E, et al. B-lymphocyte commitment: identifying the point of no return. *Semin Immunol.* 2011; 23:335–340. [PubMed: 21944938]
6. Ramirez J, et al. From hematopoietic progenitors to B cells: mechanisms of lineage restriction and commitment. *Curr Opin Immunol.* 2010; 22:177–184. [PubMed: 20207529]
7. Carpenter AC, Bosselut R. Decision checkpoints in the thymus. *Nat Immunol.* 2010; 11:666–673. [PubMed: 20644572]
8. Naito T, Taniuchi I. The network of transcription factors that underlie the CD4 versus CD8 lineage decision. *Int Immunol.* 2010; 22:791–796. [PubMed: 20732858]
9. Oestreich KJ, Weinmann AS. Master regulators or lineage-specifying? Changing views on CD4⁺ T cell transcription factors. *Nat Rev Immunol.* 2012; 12:799–804. [PubMed: 23059426]
10. O'Shea JJ, Paul WE. Mechanisms underlying lineage commitment and plasticity of helper CD4⁺ T cells. *Science.* 2010; 327:1098–1102. [PubMed: 20185720]
11. Zhu J, et al. Differentiation of effector CD4 T cell populations. *Annu Rev Immunol.* 2010; 28:445–489. [PubMed: 20192806]
12. Wirthsberger G, et al. Regulatory T-cell differentiation versus clonal deletion of autoreactive thymocytes. *Immunol Cell Biol.* 2011; 89:45–53. [PubMed: 21042335]
13. Zhou L, et al. Plasticity of CD4⁺ T cell lineage differentiation. *Immunity.* 2009; 30:646–655. [PubMed: 19464987]

14. Gavin MA, et al. Foxp3-dependent programme of regulatory T-cell differentiation. *Nature*. 2007; 445:771–775. [PubMed: 17220874]
15. Murphy KM, Stockinger B. Effector T cell plasticity: flexibility in the face of changing circumstances. *Nat Immunol*. 2010; 11:674–680. [PubMed: 20644573]
16. Mukasa R, et al. Epigenetic instability of cytokine and transcription factor gene loci underlies plasticity of the T helper 17 cell lineage. *Immunity*. 2010; 32:616–627. [PubMed: 20471290]
17. Liu X, et al. Bcl6 expression specifies the T follicular helper cell program in vivo. *J Exp Med*. 2012; 209:1841–1852. S1841–1824. [PubMed: 22987803]
18. Ostuni R, et al. Latent enhancers activated by stimulation in differentiated cells. *Cell*. 2013; 152:157–171. [PubMed: 23332752]
19. Lin YC, et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat Immunol*. 2010; 11:635–643. [PubMed: 20543837]
20. Hoogenkamp M, et al. Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. *Blood*. 2009; 114:299–309. [PubMed: 19339695]
21. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*. 2011; 25:2227–2241. [PubMed: 22056668]
22. Ostuni R, Natoli G. Lineages, cell types and functional states: a genomic view. *Curr Opin Cell Biol*. 2013
23. van Oevelen C, et al. Transcription factor-induced enhancer modulations during cell fate conversions. *Curr Opin Genet Dev*. 2013; 23:562–567. [PubMed: 23968684]
24. Vahedi G, et al. Transcription factors and CD4 T cells seeking identity: masters, minions, setters and spikers. *Immunology*. 2013; 139:294–298. [PubMed: 23586907]
25. Josefowicz SZ. Regulators of chromatin state and transcription in CD4 T-cell polarization. *Immunology*. 2013; 139:299–308. [PubMed: 23590627]
26. Mandel EM, Grosschedl R. Transcription control of early B cell differentiation. *Curr Opin Immunol*. 2010; 22:161–167. [PubMed: 20144854]
27. Laiosa CV, et al. Determinants of lymphoid-myeloid lineage diversification. *Annu Rev Immunol*. 2006; 24:705–738. [PubMed: 16551264]
28. Davidson EH. Emerging properties of animal gene regulatory networks. *Nature*. 2010; 468:911–920. [PubMed: 21164479]
29. Barolo S. Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *Bioessays*. 2012; 34:135–141. [PubMed: 22083793]
30. Hosoya-Ohmura S, et al. An NK and T cell enhancer lies 280 kilobase pairs 3' to the *gata3* structural gene. *Mol Cell Biol*. 2011; 31:1894–1904. [PubMed: 21383068]
31. Li L, et al. A far downstream enhancer for murine Bcl11b controls its T-cell specific expression. *Blood*. 2013; 122:902–911. [PubMed: 23741008]
32. Lin YC, et al. Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat Immunol*. 2012; 13:1196–1204. [PubMed: 23064439]
33. Splinter E, de Laat W. The complex transcription regulatory landscape of our genome: control in three dimensions. *EMBO J*. 2011; 30:4345–4355. [PubMed: 21952046]
34. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012; 485:376–380. [PubMed: 22495300]
35. Zhang Y, et al. Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature*. 2013; 504:306–310. [PubMed: 24213634]
36. Whyte WA, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013; 153:307–319. [PubMed: 23582322]
37. Beck D, et al. Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood*. 2013; 122:e12–22. [PubMed: 23974199]

38. Wilson NK, et al. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell*. 2010; 7:532–544. [PubMed: 20887958]
39. Ciofani M, et al. A validated regulatory network for Th17 cell specification. *Cell*. 2012; 151:289–303. [PubMed: 23021777]
40. Treiber T, et al. Early B cell Factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poising of chromatin. *Immunity*. 2010; 32:714–725. [PubMed: 20451411]
41. Wei G, et al. Genome-wide analyses of transcription factor GATA3-mediated gene regulation in distinct T cell types. *Immunity*. 2011; 35:299–311. [PubMed: 21867929]
42. Zhang JA, et al. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell*. 2012; 149:467–482. [PubMed: 22500808]
43. Heinz S, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010; 38:576–589. [PubMed: 20513432]
44. Ghisletti S, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity*. 2010; 32:317–328. [PubMed: 20206554]
45. Pham TH, et al. Mechanisms of *in vivo* binding site selection of the hematopoietic master transcription factor PU.1. *Nucleic Acids Res*. 2013; 41:6391–6402. [PubMed: 23658224]
46. McManus S, et al. The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J*. 2011; 30:2388–2404. [PubMed: 21552207]
47. Revilla-i-Domingo R, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012; 31:3130–3146. [PubMed: 22669466]
48. Lichtinger M, et al. RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J*. 2012; 31:4318–4333. [PubMed: 23064151]
49. Zarnegar MA, Rothenberg EV. Ikaros represses and activates PU.1 cell-type-specifically through the multifunctional Sfpi1 URE and a myeloid specific enhancer. *Oncogene*. 2012; 31:4647–4654. [PubMed: 22231443]
50. Chlon TM, et al. Cofactor-mediated restriction of GATA-1 chromatin occupancy coordinates lineage-specific gene expression. *Mol Cell*. 2012; 47:608–621. [PubMed: 22771118]
51. Glasmacher E, et al. A genomic regulatory element that directs assembly and function of immune-specific AP-1-IRF complexes. *Science*. 2012; 338:975–980. [PubMed: 22983707]
52. Goetz TL, et al. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor $\alpha 2$. *Mol Cell Biol*. 2000; 20:81–90. [PubMed: 10594011]
53. Gu TL, et al. Auto-inhibition and partner proteins, core-binding factor β (CBF β) and Ets-1, modulate DNA binding by CBF $\alpha 2$ (AML1). *Mol Cell Biol*. 2000; 20:91–103. [PubMed: 10594012]
54. Maier H, et al. Activation of the early B-cell-specific mb-1 (Ig- α) gene by Pax-5 is dependent on an unmethylated Ets binding site. *Mol Cell Biol*. 2003; 23:1946–1960. [PubMed: 12612069]
55. Panne D, et al. An atomic model of the interferon- β enhanceosome. *Cell*. 2007; 129:1111–1123. [PubMed: 17574024]
56. May G, et al. Dynamic analysis of gene expression and genome-wide transcription factor binding during lineage specification of multipotent progenitors. *Cell Stem Cell*. 2013; 13:754–768. [PubMed: 24120743]
57. Xu J, et al. Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. *Proc Natl Acad Sci US A*. 2007; 104:12377–12382.
58. Soufi A, et al. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*. 2012; 151:994–1004. [PubMed: 23159369]
59. Weishaupt H, et al. Epigenetic chromatin states uniquely define the developmental plasticity of murine hematopoietic stem cells. *Blood*. 2010; 115:247–256. [PubMed: 19887676]
60. Vigano MA, et al. An epigenetic profile of early T-cell development from multipotent progenitors to committed T-cell descendants. *Eur J Immunol*. 2013

61. Miller SA, et al. Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol Cell*. 2010; 40:594–605. [PubMed: 21095589]
62. Koche RP, et al. Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell*. 2011; 8:96–105. [PubMed: 21211784]
63. Zhang B, et al. A dynamic H3K27ac signature identifies VEGFA-stimulated endothelial enhancers and requires EP300 activity. *Genome Res*. 2013; 23:917–927. [PubMed: 23547170]
64. Mercer EM, et al. Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity*. 2011; 35:413–425. [PubMed: 21903424]
65. Rada-Iglesias A, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011; 470:279–283. [PubMed: 21160473]
66. Creighton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010; 107:21931–21936. [PubMed: 21106759]
67. Garber M, et al. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol Cell*. 2012; 47:810–822. [PubMed: 22940246]
68. Sekiya T, Zaret KS. Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. *Mol Cell*. 2007; 28:291–303. [PubMed: 17964267]
69. Xie W, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell*. 2013; 153:1134–1148. [PubMed: 23664764]
70. Ku M, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet*. 2008; 4:e1000242. [PubMed: 18974828]
71. Vahedi G, et al. STATs shape the active enhancer landscape of T cell populations. *Cell*. 2012; 151:981–993. [PubMed: 23178119]
72. Zheng Y, Rudensky AY. Foxp3 in control of the regulatory T cell lineage. *Nat Immunol*. 2007; 8:457–462. [PubMed: 17440451]
73. Ziegler SF. FOXP3: of mice and men. *Annu Rev Immunol*. 2006; 24:209–226. [PubMed: 16551248]
74. Samstein RM, et al. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell*. 2012; 151:153–166. [PubMed: 23021222]
75. Mercer TR, et al. DNase I-hypersensitive exons colocalize with promoters and distal regulatory elements. *Nat Genet*. 2013; 45:852–859. [PubMed: 23793028]
76. Neph S, et al. Circuitry and dynamics of human transcription factor regulatory networks. *Cell*. 2012; 150:1274–1286. [PubMed: 22959076]
77. Ji H, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature*. 2010; 467:338–342. [PubMed: 20720541]
78. Pekowska A, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J*. 2011; 30:4198–4210. [PubMed: 21847099]
79. Mingueneau M, et al. The transcriptional landscape of $\alpha\beta$ T cell differentiation. *Nat Immunol*. 2013; 14:619–632. [PubMed: 23644507]
80. David-Fung ES, et al. Transcription factor expression dynamics of early T-lymphocyte specification and commitment. *Dev Biol*. 2009; 325:444–467. [PubMed: 19013443]
81. Kawazu M, et al. Expression profiling of immature thymocytes revealed a novel homeobox gene that regulates double-negative thymocyte development. *J Immunol*. 2007; 179:5335–5345. [PubMed: 17911620]
82. Wei G, et al. Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4⁺ T cells. *Immunity*. 2009; 30:155–167. [PubMed: 19144320]
83. Rothenberg EV, Zhang JA. T-cell identity and epigenetic memory. *Curr Top Microbiol Immunol*. 2012; 356:117–143. [PubMed: 21833836]
84. Miyazaki KMM, Murre C. Establishment of B versus T cell identity. *Trends Immunol*. 2014 in press.

85. Hosoya T, et al. From the cradle to the grave: activities of GATA-3 throughout T-cell development and differentiation. *Immunol Rev.* 2010; 238:110–125. [PubMed: 20969588]
86. Wang Y, et al. GATA-3 controls the maintenance and proliferation of T cells downstream of TCR and cytokine signaling. *Nat Immunol.* 2013; 14:714–722. [PubMed: 23708251]
87. Wang R, et al. T cell factor 1 regulates thymocyte survival via a ROR γ t-dependent pathway. *J Immunol.* 2011; 187:5964–5973. [PubMed: 22039299]
88. Månsson R, et al. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity.* 2007; 26:407–419. [PubMed: 17433729]
89. Ng SY, et al. Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. *Immunity.* 2009; 30:493–507. [PubMed: 19345118]
90. Hu M, et al. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev.* 1997; 11:774–785. [PubMed: 9087431]
91. Miyamoto T, et al. Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev Cell.* 2002; 3:137–147. [PubMed: 12110174]
92. Moignard V, et al. Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat Cell Biol.* 2013; 15:363–372. [PubMed: 23524953]
93. Pina C, et al. Inferring rules of lineage commitment in haematopoiesis. *Nat Cell Biol.* 2012; 14:287–294. [PubMed: 22344032]
94. Guo G, et al. Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. *Cell Stem Cell.* 2013; 13:492–505. [PubMed: 24035353]
95. Satija RS, AK. Heterogeneity in immune responses -- from populations to single cells. *Trends Immunol.* 2014 in press.
96. Rodriguez-Ubreva J, et al. Pre-B cell to macrophage transdifferentiation without significant promoter DNA methylation changes. *Nucleic Acids Res.* 2012; 40:1954–1968. [PubMed: 22086955]
97. Egawa T, Littman DR. ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. *Nat Immunol.* 2008; 9:1131–1139. [PubMed: 18776905]
98. Zullo JM, et al. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell.* 2012; 149:1474–1487. [PubMed: 22726435]

Highlights

- Multi-tiered action of transcription factors remodels the chromatin landscape
- Differential chromatin access and partner factors bias binding site choice
- Powerful transcription factors can work by modifying effects of other factors
- Selective transcription factor silencing shapes the regulatory state in early T cells

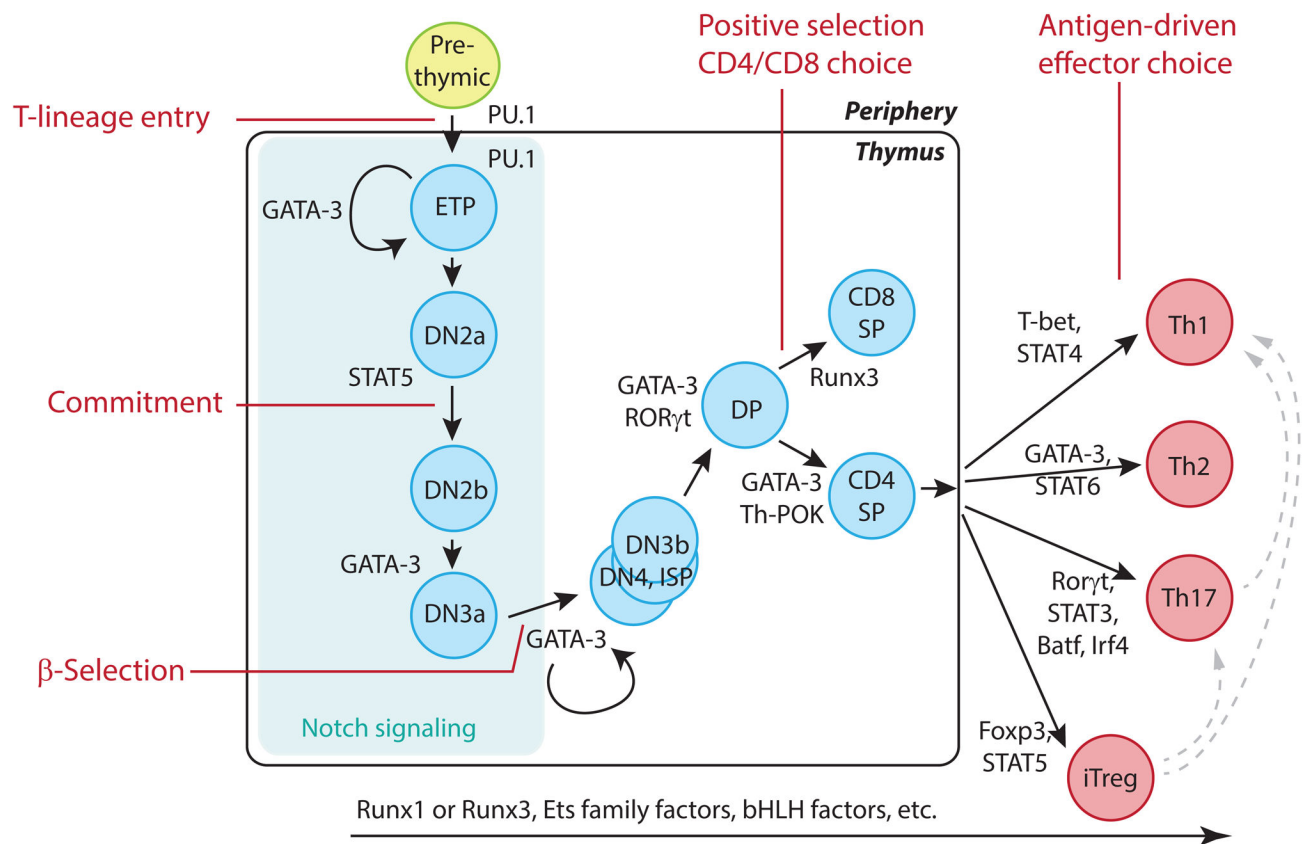


Figure 1. Decision points and transcription factors in T cell development

A simplified schematic of T-cell development is shown with the main stages of development within the thymus labeled, key checkpoints and choice points indicated in red labels, and selected transcription factors indicated at major points when they must act. The figure shows diagrammatically which events occur in the thymus and which ones occur either prethymically or in the periphery after T-cell maturation. Major decisions and stages that depend on GATA-3, PU.1, ROR γ t, T-bet, Foxp3, and multiple STAT factors are indicated. Runx, Ets, and bHLH family factors have continuing roles throughout these events, although the precise family members involved and their roles can change. T-lineage entry: migration of cells to the thymic microenvironment, which is rich in Notch-activating molecules. Commitment: relinquishment of potential to give rise to anything except T cells. β -selection, positive selection: two stages of TCR-dependent developmental checkpoint where cells with failed TCR status will die. Also shown are the stages when continuing Notch pathway signaling in the thymus is critical for T-cell specification: these are stages when the cells are “double negative”, i.e. lacking CD4 and CD8 expression. The stages within this immature series are ETP (Early T-cell Precursor), c-Kit⁺ CD44⁺ CD25⁻; DN2a, c-Kit⁺ CD44⁺ CD25⁺; DN2b, c-Kit^{+(intermediate)} CD44⁺ CD25⁺; DN3a, c-Kit⁻ CD44⁻ CD25⁺. TCR gene rearrangement does not result in expression of a TCR β chain until the DN3a stage, but this triggers β -selection. The cells quickly downregulate CD25 and upregulate CD8 and CD4 through the series of stages shown as DN3b, DN4, ISP in transition to “DP”, or CD4⁺ CD8⁺ double-positive. These cells continue TCR gene rearrangement until they finally express

TCR $\alpha\beta$. They are subject to stringent TCR-based MHC restriction and selection and only a minority are allowed to mature as CD4 SP or CD8 SP cells (Positive selection, CD4/CD8 choice). For simplicity, alternative intrathymic fates such as TCR $\gamma\delta$ cells, invariant NKT cells (iNKT) or natural Tregs (natural Tregs) are not shown. Four major types of effector subsets of CD4⁺ cells, which develop from mature CD4 SP cells in the periphery, are shown, with some of the transcription factors that play major roles in their development. Despite the clear distinctions between the regulatory programs of these effector subsets, they remain conditionally plastic depending on environmental signals. Gray dashed-line arrows indicate several well-established options for interconversion. This feature contrasts the effector cell lineage decisions with the decision to commit to T-cell fate, which is irreversible.

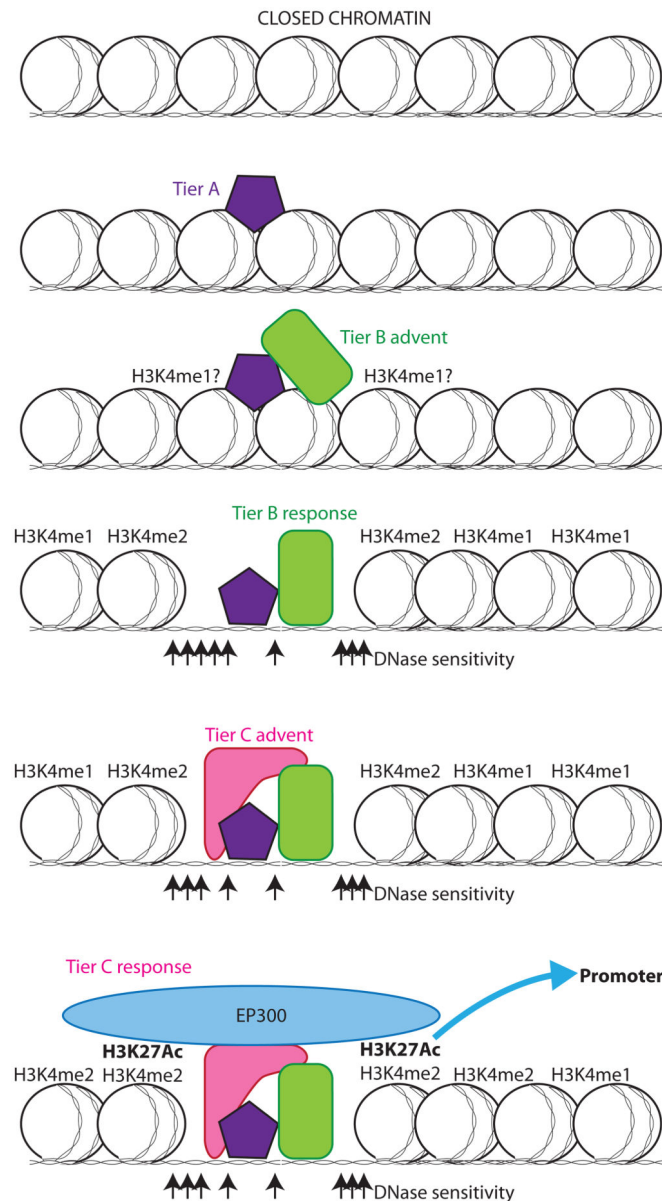


Figure 2. Multiple tiers of transcription factor action generate an active enhancer

The figure shows three kinds of transcription factor occupancy that can have distinct effects on the activity of an enhancer element. The steps in converting a closed, silent cis-regulatory region to an active enhancer are shown for an element in which three factors (or groups of factors) act in sequence. Only the Tier A factors are capable of recognizing their sites in the DNA from a compact chromatin configuration. Whether or not the Tier A factors themselves recruit the enzymes that mark local histones as “accessible”, however, they make it easier for Tier B factors to bind. Tier B factors are those that elicit activating histone marks and displace nucleosomes to create a DNase-accessible site. Note that the effect of Tier B occupancy is to create a small “volcano”-like shape of histone modification consisting of a wide, hollowed-out H3K4me1 peak with a narrow, cratered peak of H3K4me2 in its center. Still, despite this modification, the cis-regulatory element may not

be active yet. Then, as Tier C factors arrive and bind via DNA and/or protein-protein interactions, a quorum is finally complete for the recruitment of the histone acetyltransferase EP300. Tier C factors may bind DNA directly or, as in the case of Foxp3, can determine the functional output of an enhancer even without extensive DNA contact. The large EP300 protein catalyzes modification of proximal histones to H3K27Ac forms and completes the activation of the cis-regulatory element. Whether or not the EP300 remains associated at the site, these fully activated, occupied elements are now competent to interact with the appropriate transcription start site to enhance transcription.